

실시간 차량 밀도에 대응하는 심층강화학습 기반 C-V2X 분산혼잡제어

Deep Reinforcement Learning-Based C-V2X Distributed Congestion Control for Real-Time Vehicle Density Response

전 병 철*, 양 우 열*, 조 한 신**★

Byeong Cheol Jeon*, Woo Yoel Yang*, Han-Shin Jo**★

Abstract

Distributed congestion control (DCC) is a technology that mitigates channel congestion and improves communication performance in high-density vehicular networks. Traditional DCC techniques operate to reduce channel congestion without considering quality of service (QoS) requirements. Such design of DCC algorithms can lead to excessive DCC actions, potentially degrading other aspects of QoS. To address this issue, we propose a deep reinforcement learning-based QoS-adaptive DCC algorithm. The simulation was conducted using a quasi-real environment simulator, generating dynamic vehicular densities for evaluation. The simulation results indicate that our proposed DCC algorithm achieves results closer to the targeted QoS compared to existing DCC algorithms.

요 약

분산혼잡제어는 높은 밀도의 차량 네트워크에서 채널 혼잡을 완화하고, 통신 성능을 개선하는 기술이다. 기존 분산혼잡제어 기술은 quality of service(QoS) 요구사항을 고려하지 않은 채 채널 혼잡을 줄이는 방향으로 동작한다. 이러한 분산혼잡제어 알고리즘 설계는 과도한 DCC 동작으로 인하여 다른 QoS를 저하시킬 수 있다. 이와 같은 문제를 해결하기 위해 심층강화학습 기반 QoS 적응형 DCC 알고리즘을 제안한다. 시뮬레이션은 준 실험 환경 시뮬레이터를 기반으로 동적인 차량 밀도를 생성하여 평가하였으며, 시뮬레이션 결과 기존 DCC 알고리즘 보다 목표 QoS에 더 근접한 결과를 확인하였다.

Key words : Cellular-V2X, Distributed Congestion Control, Deep Reinforcement Learning, Quality of Service, Microscopic Simulator

* Dept. of Electronic Engineering, Hanbat National University

** Dept. of Automotive Engineering, Hanyang University

★ Corresponding author

E-mail : hsjo23@hanyang.ac.kr, Tel : 02-2220-2891

※ Acknowledgment

This work was supported in part by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2022-0-01053)

Manuscript received Nov. 23, 2023; revised Dec. 4, 2023; accepted Dec. 8, 2023.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

자율주행 자동차는 카메라, LiDAR, RADAR 등 다양한 센서 및 전장부품을 통해 주변 환경을 인식하여 운전자의 개입 없이 스스로 주행 상황을 인지하고 자동차를 제어한다. 미국의 국제자동차기술자협회(Society of Automotive Engineers, SAE)는 이러한 자율주행을 자동화 수준에 따라 표 1.1과 같이 6단계로 분류하고 있다[1]. 완전 자율주행 수준인 레벨 4는 운전자의 개입 없이 서비스를 지원해야 하며, 자동차의 전장 부품만으로 해당 서비스를 지원하기에는 한계가 있다. 이러한 한계를 극복하기 위해 차세대 지능형 교통 시스템 (Cooperative-Intelligent Transport Systems, C-ITS)와의 통합이 중요하다. C-ITS는 주행 중 차량이 운전자에게 주변의 교통상황, 전방의

장애물 등 사고 위험 정보를 실시간으로 제공한다. 이와 같은 C-ITS를 구축하기 위한 필수 기술로는 V2X (Vehicle-to-Everything)가 있다. 해당 기술은 차량과 차량(Vehicle to Vehicle, V2V), 차량과 보행자(Vehicle to Pedestrian, V2P), 차량과 인프라(Vehicle to Infrastructure, V2I) 등 유·무선망을 통해 다른 차량 및 도로 등 인프라가 구축된 사물과 정보를 교환하는 기술을 총칭한다.

C-V2X(Cellular-V2X)는 이동 통신 표준화 기관인 3GPP(3rd Generation Partnership Project)의 LTE (Long-Term Evolution) 통신 프로토콜 기반의 기술로 C-V2X의 V2V 시나리오는 CAM(Cooperative Awareness Message)를 통해 차량의 위치, 속도 그리고 방향 등의 정보를 주기적으로 주변 차량과 정보를 교환한다.

3GPP에서 [1]의 자동화 수준에 따라 군집 주행, 첨단 주행, 센서 확장, 원격 주행 4가지 서비스에 대한 서로 다른 QoS(Quality of Service) 요구사항을 정의한다 [2], [3]. C-V2X의 CAM은 4가지 서비스를 지원하기 위해 필수적이다.

차량 밀도의 증가는 한정된 무선 자원을 다수의 차량이 점유하기 때문에 패킷 충돌을 야기해 성능 저하를 일으킨다. 분산 혼잡 제어(Distributed Congestion Control, DCC)는 이러한 고밀도 환경의 통신 성능 개선을 위한 대표적인 기술이다. DCC는 메시지 전송 주기, 전송 전력, modulation and coding scheme(MCS) 등을 조절하여 채널 혼잡도를 낮추고, 통신 성능을 개선한다. 하지만 기존 차량 통신 관련 표준화 기관에서 제안하는 DCC 기술은 채널 혼잡도를 감소시키는 것에 초점을 맞추어 통신 성능을 개선하기 때문에 서비스의 QoS 요구사항을 만족할 수 있는 환경이 제한적이다.

따라서, 본 논문에서는 심층강화학습 기반의 QoS 요구사항에 적응적인 DCC 알고리즘을 제안한다. 심층강화학습은 강화학습과 딥러닝을 결합한 기계 학습의 하위 분야로 V2X 분야에도 다양하게 연구되고 있다[4], [5]. 또한 C-V2X V2V 시나리오에서 동작 가능한 심층강화학습 모델을 제안하고, 이를 준 실환경 시뮬레이터를 활용하여 동적 차량 밀도를 모델링하여 이 알고리즘을 적용한다. 이를 통해 실제 C-V2X 시스템 구축을 위한 알고리즘의 가능성을 검증한다.

II. 시스템 모델

C-V2X는 single-carrier frequency-division multiple

access(SC-FDMA)을 활용하며, 10 MHz 및 20 MHz의 채널 대역폭에서 운영된다. 채널은 시간 및 주파수 영역에서 각각 서브프레임과 서브채널로 구분된다. 서브프레임의 길이는 1ms이고, 서브채널은 사용자에게 할당되는 가장 작은 자원의 단위인 resource block(RB)로 구성된다. RB는 시간 영역에서 0.5 ms, 주파수 영역에서 180 kHz를 차지한다, 패킷은 1 ms의 서브프레임 내에서 1 개 혹은 다수의 서브채널을 통해 브로드캐스트된다. V 대의 차량 사용자 단말(vehicle user equipment, VUE)은 주어진 메시지 전송 주기 내의 n 번째 서브프레임의 주파수 영역에서 z 번째 서브채널을 통해 CAM (cooperative awareness message)를 전송한다. 여기서 n 은 $\mathbf{N} = \{n | n = 1, 2, \dots, N\}$ 로 정의되는 서브프레임의 인덱스이며, z 는 $\mathbf{Z} = \{z | z = 1, 2, \dots, Z\}$ 로 정의되는 n 번째 서브프레임에서 사용 가능한 서브채널의 집합이다. 그리고 V 대의 VUE는 $\mathbf{V} = \{v | v = 1, 2, \dots, V\}$ 로 정의되며, 독립적으로 sensing based semi-persistent scheduling(SB-SPS)를 통해 무선 자원을 할당한다[6].

전송된 CAM의 수신 여부는 signal-to-interference-plus-noise ratio (SINR)에 의해 결정된다. v 번째 VUE가 n 번째 서브프레임의 z 번째 서브채널을 통해 패킷을 브로드캐스트하는 경우 l 번째 VUE가 수신하는 SINR은 다음과 같이 계산된다.

$$\gamma_{v,l}[n,z] = \frac{S_{v,l}[n,z]}{\sigma^2 + I_{v,l}[n,z]} \quad (1)$$

여기서 σ^2 는 잡음 전력이며, l 번째 VUE가 v 번째 VUE로부터 수신하는 전력은 다음과 같이 계산된다.

$$S_{v,l}[n,z] = P_{tx} G_{tx} G_{rx} \eta_{v,l} \quad (2)$$

여기서 P_{tx} 는 송신 전력, G_{rx} 와 G_{tx} 는 각각 수신 안테나 이득, 송신 안테나 이득을 의미한다. $\eta_{v,l}$ 은 경로 손실과 음영을 고려한 대규모 페이딩이다. v 번째 VUE와 l 번째 VUE가 동일 서브채널에 무선 자원을 점유한 경우 누적 간섭 전력은 다음과 같이 계산된다.

$$I_{v,l}[n,z] = \sum_{q \in \mathbf{Q}, q \neq v} P_{tx} G_{tx} G_{rx} \eta_{v,l} \quad (3)$$

\mathbf{Q} 는 v 번째 VUE와 동일한 무선 자원을 점유하는 VUE의 집합이다. 식 (1)을 통해 측정된 $\gamma_{v,l}[n,z]$ 가 메시지의 크기와 MCS에 의해 결정되는 요구되는 최소 SINR γ_{\min} 보다 클 경우 패킷을 올바르게 수신하고, 그

렇지 않으면 수신되지 않는다.

QoS 요구사항 중 신뢰도를 의미하는 packet delivery ratio (PDR)은 총 브로드캐스트된 패킷의 수와 수신된 패킷 수의 비율로써 정의된다. CAM은 통신 범위 β 내 모든 VUE에 브로드캐스트되며, n 번째 서브프레임에서 v 번째 VUE가 브로드캐스트한 패킷의 PDR은 다음과 같이 계산된다.

$$\theta_v[n, z, \beta] = \frac{\delta_v[n, z, \beta]}{\lambda_v[n, z, \beta]} \quad (4)$$

여기서 $\delta_v[n, z, \beta]$ 는 통신 범위 β 내에서 올바르게 수신된 패킷의 수를 의미하며, $\lambda_v[n, z, \beta]$ 는 총 브로드캐스트된 패킷의 수를 의미한다.

III. 심층강화학습 기반 QoS 적응형 전송 주기 제어 알고리즘

1. 최적화 문제 정의

VUE는 메시지 전송 주기 제어를 통해 통신 범위 내 PDR을 조절할 수 있다. 또한 전송 주기 제어의 동작을 통해 차량 네트워크의 혼잡도를 낮추어 PDR을 향상시킬 수 있다. 하지만 혼잡 수준에 대한 과도한 전송 주기 제어의 동작은 inter packet gap(IPG)를 증가시킬 수 있다. 따라서 혼잡한 상황에서 적절한 수준의 메시지 전송 주기 제어 동작은 필수적이다. 따라서 본 논문에서 VUE는 목표 통신 범위 내에서 목표 PDR을 만족하는 수준까지만 전송 주기 제어를 동작시킨다. 최적화 문제는 다음과 같이 정의된다.

$$P1 : \min_{\{U\}} |\psi_{\rho, \beta}^* - \psi_{\rho, \beta}^{avg}|, \quad (5)$$

$$\psi_{\rho, \beta}^{avg} = \frac{\sum_{v \in \mathbf{V}} \sum_{z \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \delta[n, c, \beta]}{\sum_{v \in \mathbf{V}} \sum_{z \in \mathbf{Z}} \sum_{n \in \mathbf{N}} \lambda[n, c, \beta]},$$

$$s.t. C1 : \mathbf{U} = \{u_1, u_2, u_3, \dots, u_M\}$$

$$u_m \in \{\delta_1, \delta_2, \delta_3, \dots, \delta_F\}$$

식 (5)에서 $\psi_{\rho, \beta}^*$ 는 차량 밀도 ρ 에 대해 목표 통신 범위 β 에서의 차량 네트워크 목표 PDR이며, $\psi_{\rho, \beta}^{avg}$ 는 차량 네트워크 평균 PDR이다. u_m 은 패킷 생성 시간 간격의 상한값을 의미하며, 이는 δ 집합의 원소이다. 패킷 생성 시간은 최소 100 ms와 u_m 사이의 무작위 정수로 결정된다. \mathbf{U} 는 시나리오 내 모든 차량의 패킷 생성 시간 간격

의 집합이다. 결과적으로 $P1$ 은 차량 네트워크 상 모든 VUE의 패킷 생성 시간을 분산적으로 조절하여 통신 범위 β 내 평균 PDR과 목표 PDR의 차이를 최소화하는 문제로 정의된다.

2. Deep Q-Network(DQN)

강화학습은 학습하는 주체인 에이전트가 환경과 상호 작용하여 보상을 최대화하는 최적의 행동 또는 전략을 학습하는 과정이다[7]. 심층강화학습은 딥러닝의 신경망을 도입하여 정책 혹은 행동 가치 함수를 추정할 수 있다. DQN은 심층강화학습 기반의 value-based 알고리즘 중 가장 대표적인 알고리즘이다[8]. DQN의 핵심 아이디어는 experience replay와 두 개의 신경망을 사용하는 학습 방법이다. experience replay는 각 시간 단계에서의 경험(s_t, a_t, r_t, s_{t+1})을 D 용량의 replay memory에 저장한다. 학습 과정에서 이 memory에서 무작위의 샘플을 추출하여 학습한다. 이와 같은 학습 방법은 데이터 간의 상관 관계를 줄이고 학습의 안정성을 향상시킨다. 또한 두 개의 신경망 로컬 네트워크 $Q(s, a; \theta)$ 와 타겟 네트워크 $Q(s, a; \theta^-)$ 는 Q 값의 발산 문제를 완화시킬 수 있는 장점이 있다. 로컬 네트워크는 $\epsilon - greedy$ 정책에 따라 $1 - \epsilon$ 의 확률로 행동을 무작위로 선택하고, ϵ 의 확률로 다음과 같이 행동을 선택한다.

$$a_t = \arg \max_a Q(s_t, a; \theta) \quad (6)$$

식 (6)에서 로컬 네트워크는 현재 상태 s_t 에서 가장 큰 Q 값을 가질 수 있는 행동 a_t 를 선택한다. 행동에 따라 환경은 에이전트에게 행동에 따른 보상 r_t 와 다음 상태 s_{t+1} 에 대한 정보를 제공한다. 환경으로부터 제공받은 transition(s_t, a_t, r_t, s_{t+1})은 replay memory에 저장된다. transition이 사전 구성된 미니 배치 사이즈 J 보다 큰 경우 에이전트는 $1 \times J$ 개의 무작위 transition 샘플을 추출하여 로컬 네트워크를 업데이트한다. 로컬 네트워크의 신경망 θ 를 업데이트 하기 위한 손실함수는 다음과 같이 정의된다.

$$L(\theta) = E[(y_j - Q(s_j, a_j; \theta))^2] \quad (7)$$

여기서 y_j 는 목표 값으로 다음과 같이 정의된다.

$$y_j = \begin{cases} r_j & s_{j+1} = s_T \\ r_j + \gamma \max_{a' \in A} Q(s_{j+1}, a'; \theta^-) & s_{j+1} \neq s_T \end{cases} \quad (8)$$

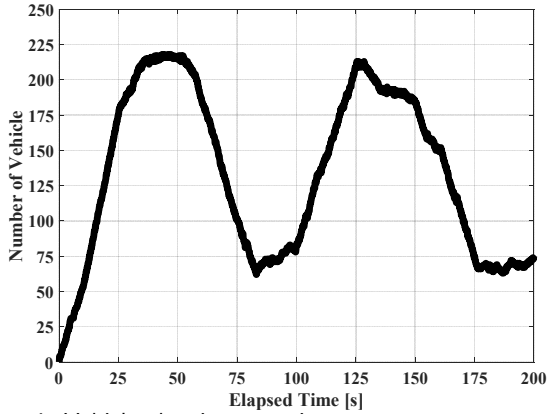


Fig. 1. Vehicle density over time.
그림 1. 시간 경과에 따른 차량 수

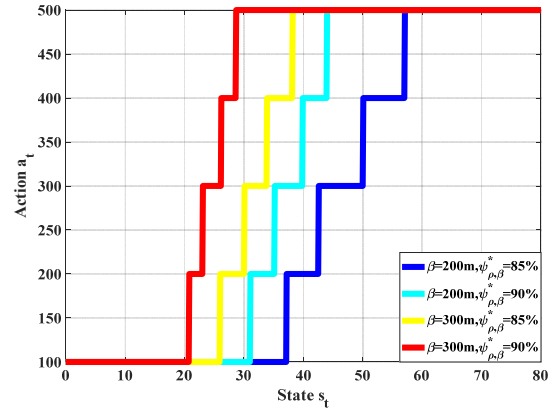
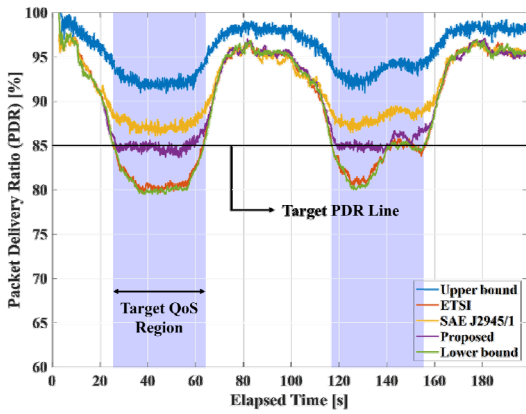
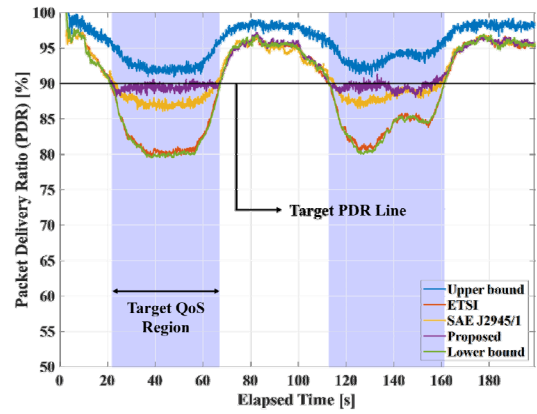


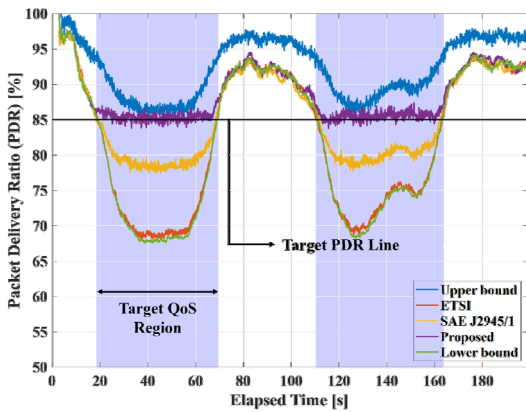
Fig. 2. Message rate control policies.
그림 2. 메시지 전송 주기 정책



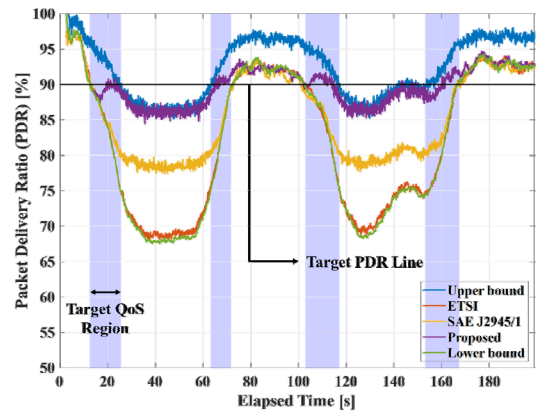
(a) $\beta=200$ m, $\psi_{\rho,\beta}^*=85\%$



(b) $\beta=200$ m, $\psi_{\rho,\beta}^*=90\%$



(c) $\beta=300$ m, $\psi_{\rho,\beta}^*=85\%$



(d) $\beta=300$ m, $\psi_{\rho,\beta}^*=90\%$

Fig. 3. Vehicle network instantaneous PDR over time.
그림 3. 시간 경과에 따른 차량 네트워크 순시적 PDR

에이전트는 위와 같은 과정을 통해 네트워크를 학습시켜 설계된 보상을 최대화할 수 있는 행동을 선택한다.

3. 심층강화학습 모델

다수의 VUE가 공존하는 V2V 시나리오에서 멀티 에이전트 심층강화학습은 효과적인 학습 프레임워크 중 하나

로 간주된다. 해당 학습 프레임워크는 동시에 학습하는 환경이기 때문에 각 에이전트의 행동이 환경에 영향을 주어 다른 에이전트의 학습에도 영향을 줄 수 있다. 이러한 MARL 학습 환경은 비정상성으로 인하여 markov property를 붕괴시킨다. 비정상성 문제를 해결하기 위한 MARL 학습 프레임워크로는 fully centralized,

centralized training with decentralized execution (CTDE), networked agents 등이 있다. 그중 CTDE는 훈련과 실행 단계로 구분되며, 훈련 단계에서 중앙 집중식 컨트롤러는 모든 에이전트의 정보를 수집하여 각 에이전트를 위한 단일 정책을 학습한다. 학습된 단일 정책은 실행 단계에서 모든 에이전트에게 공유된다. 따라서 실행 단계에서 에이전트는 중앙 집중식 컨트롤러로부터 전역 정보를 제공받지 않아도 된다. 따라서 본 논문에서는 V2V mode 4 시나리오에서 전역 정보를 활용하지 않고 지역 정보만을 활용할 수 있는 CTDE 학습 방식을 채택했다. CTDE의 학습 단계에서 중앙 집중식 컨트롤러는 roadside unit(RSU)로 대체할 수 있으며, 학습을 위한 상태는 다음과 같다.

$$s_t = \frac{1}{M} \sum_{m \in M} \eta_t^{(m)} \quad (9)$$

여기서 $\eta_t^{(m)}$ 는 시간 t 에서 m 번째 차량의 100 m 이내의 차량 수를 의미한다. 즉, 각 차량마다 측정된 100 m 이내의 차량수의 평균이다. 이와같이 차량 수를 활용하여 채널의 혼잡을 추정할 수 있는 지표는 SAE J2945/1 전송 주기 제어 알고리즘에서도 활용된다. 컨트롤러는 상태 정보를 기반으로 목표 QoS에 근접하도록 전송 주기를 조절해야 한다. 컨트롤러의 행동 공간은 다음과 같다.

$$A = \{\delta_1, \delta_2, \delta_3, \dots, \delta_F\}, a_t \in A \quad (10)$$

여기서 $F=5$ 이며, $a_t = \{100, 200, 300, 400, 500\}$ 으로 설정된다. 학습의 목표인 $\psi_{\rho, \beta}^*$ 와 $\psi_{\rho, \beta}^{avg}$ 사이를 최소화하기 위한 학습을 진행한다. 각 차량은 전역 정보인 $\psi_{\rho, \beta}^{avg}$ 를 알 수 없지만, 중앙 컨트롤러는 전역 정보를 안다고 가정한다. 학습의 보상 설계는 다음과 같다.

$$r_t = -|\psi_{\rho, \beta}^* - \psi_{\rho, \beta}^{avg}| \quad (11)$$

학습이 종료된 뒤, 각 차량은 학습된 정책을 컨트롤러로부터 공유받으며 실행 단계에서 컨트롤러의 개입은 요구되지 않는다. 실행 단계에서의 각 차량은 로컬 정보 $\eta_t^{(m)}$ 만을 이용하여 독립적으로 동작한다.

4. 성능 평가

가. 벤치마크 알고리즘

제안하는 알고리즘을 평가하기 위해 ETSI의 DCC 알

고리즘과 SAE J2945/1 DCC 알고리즘을 벤치마크 알고리즘으로 채택한다.

(1) ETSI DCC 프레임워크

ETSI는 채널의 혼잡을 제어하기 위해 channel busy ratio(CBR)과 channel occupancy ratio(CR) 두 가지 지표를 활용한다[9]. 각 VUE는 측정된 CBR을 기반으로 CR limit를 결정하며, CR이 CR limit보다 높은 경우 메시지의 전송 주기를 조절하여 CR을 CR limit 아래로 줄일 수 있다. 표 1은 시뮬레이션에 사용된 측정된 CBR 기반 CR limit 값을 보여준다.

(2) SAE J2945/1

SAE는 DSRC 기반의 DCC 알고리즘을 제안한다 [10]. 해당 알고리즘은 메시지의 전송 주기와 전송 전력을 동시 제어한다. 전송 주기는 100 m 이내의 노드 수 (vehicle density, VD)에 따라 inter-transmission time(ITT)를 조정하며 다음과 같이 계산된다.

$$T = \begin{cases} 0.1 & (VD \leq 25) \\ \frac{VD}{250} & (25 < VD < 150) \\ 0.6 & (150 \leq VD) \end{cases} \quad (12)$$

전송 전력은 측정된 CBR을 기반으로 제어되며 다음과 같이 계산된다.

$$P_{tx} = \begin{cases} 23 & (CBR \leq 50) \\ 23 - \frac{13}{30}(CBR - 50) & (50 < CBR < 80) \\ 10 & (80 \leq CBR) \end{cases} \quad (13)$$

각 제어 변수는 차량 네트워크상 고밀도 환경에서 채널 혼잡도를 감소시키는 방향으로 동작한다.

나. 시뮬레이션 결과

제안하는 알고리즘을 평가하기 위한 시뮬레이션 파라미터는 표 1과 같다. 알고리즘에서 목표 QoS는 통신 범위 β 와 PDR로 채택했으며, 각각 $\beta=200, 300$ m 와 $\psi_{\rho, \beta}^*=85, 90\%$ 이다. 실행 단계의 도로 환경 및 차량의 이동성은 microscopic 시뮬레이터 PTV VISSIM을 통해 모델링하였으며, 그림 1과 같이 시간에 따라 시스템 상 차량의 수가 동적으로 변한다. 이를 통해 C-V2X 시스템의 실제 환경에서의 적용 가능성을 확인한다.

그림 2는 학습된 에이전트의 행동 정책을 보여준다. 학습된 행동 정책에서 입력 상태가 클수록 에이전트는 높은 행동을 취한다. 또한 높은 수준의 목표 QoS가 요

구될수록 낮은 상태 입력 값에서 더 큰 행동을 선택한다.

그림 3은 시간 경과에 따른 차량 네트워크의 순시적 PDR을 나타낸다. 이 결과에서 Lower Bound와 Upper Bound는 행동 공간의 하한과 상한을 대변한다. (a) 섹션의 Target QoS Region에서 SAE J2945/1의 평균 PDR은 87.61%로 나타나 하한 대비 약 6%의 성능 향상을 보였다. 반면, ETSI는 82.21%의 PDR을 보여 단지 0.5%의 미약한 향상을 나타냈다. 제안된 알고리즘은 약 84.98%의 평균 PDR을 보이며 목표 QoS($\psi_{\rho, \beta}^* = 85\%$)에 가장 근접한 성능을 보였다. (b) 섹션에서는 SAE J2945/1, ETSI, 그리고 제안된 알고리즘의 평균 PDR이 각각 88.04%, 83.26%, 89.29%로 측정되었다. ETSI는 (a)의 결과와 유사한 경향을 보이는 반면, 제안된 알고리즘은 목표 QoS($\psi_{\rho, \beta}^* = 90\%$)에 가장 근접한 결과를 나타냈다.

Table 1. Simulation parameter.

표 1. 시뮬레이션 파라미터

Parameters	Value
Simulation time	200 s
Message size	300 Bytes
Channel bandwidth	10 MHz
MCS index	3
Channel model	WINNER + B1
Antenna gain (G_{tx} , G_{rx})	3 dB
Noise figure	9 dB
Target comm. range (β)	200, 300 m
Target PDR ($\psi_{\rho, \beta}^*$)	85, 90%
Memory capacity (D)	100000
Mini-batch size (J)	1024

(c) 및 (d) 섹션에서도 (b)와 유사한 경향이 관찰된다. 목표 QoS가 85% 인 경우 제안된 알고리즘은 상한 및 하한 내에서 평균 84.93%의 평균 PDR을 보이며, ETSI와 SAE J2945/1은 각각 73.30%, 79.79%의 평균 PDR을 기록했다. 90%의 목표 QoS에서는 제안된 알고리즘이 89.22%의 평균 PDR을 나타내며, ETSI와 SAE J2945/1은 각각 81.76%, 83.94%로 측정되었다.

결론적으로 벤치마크로 활용된 두 알고리즘은 향상된 평균 PDR을 보이나 주어진 목표 QoS를 만족시키지 못했다. 반면, 제안된 알고리즘은 더 효과적인 결과를 보여줌으로써 차량 네트워크에서의 성능 향상에 기여할 수 있음을 시사한다.

IV. 결론

본 연구는 심층강화학습 기반의 QoS 요구사항 적응형 DCC 알고리즘을 제안한다. 제안하는 알고리즘은 기존 DCC에서 고려하지 않았던 QoS 요구사항을 활용함으로써 목표 QoS 수준을 효과적으로 만족시켰다. 준 실험 환경 시뮬레이터의 동적 차량 밀도를 모델링한 시뮬레이션 결과를 통해 제안하는 알고리즘이 기존 DCC보다 QoS를 더 잘 만족하는 것을 입증하였다. 이러한 과정을 통하여 C-ITS를 구축하기 위한 C-V2X 시스템의 실제 적용 가능성을 확인하였다. 향후 연구에서는 제안하는 알고리즘 및 학습 모델을 도심과 같이 다양한 토폴로지 및 전파 환경 특성을 고려한 보다 일반화된 심층강화학습 모델로 확장하는 것을 고려한다.

References

- [1] *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, Standard SAE J3016, Apr. 2021.
- [2] *Technical Specification Group Services and System Aspects; Study on enhancement of 3GPP Support for 5G V2X Services*, document TR 22.886 V15.3.0, 3GPP Sep. 2018.
- [3] *Enhancement of 3GPP support for V2X scenarios*, document TS 22.186 v15.3.0, 3GPP, Jun. 2018.
- [4] J. Choi, H. Jo, C. Mun, and J. Yook, "Deep Reinforcement Learning-Based Distributed Congestion Control in Cellular V2X Networks," *IEEE Wireless Commun. Lett.*, vol.10, no.11, pp.2582-2586, Nov. 2021. DOI: 10.1109/LWC.2021.3108821
- [5] M. Roshdi, S. Bhadauria, K. Hassan, and G. Fischer, "Deep reinforcement learning based congestion control for V2X communication," in *Proc. IEEE 32nd Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2021, pp. 1-6. DOI: 10.1109/PIMRC50174.2021.9569259
- [6] *Evolved Universal Terrestrial Radio Access (E-UTRA)*, document TS 36.213 v 14.8.0, 3GPP, Oct. 2018.
- [7] R. S. Sutton and A. G. Barto, *Reinforcement*

Learning: An Introduction. Cambridge, MA, USA: MIT Press, 2018.

[8] V. Mnih et al., "Human-level control through deep reinforcement learning" in *Nature*, vol.518, pp.529-533, 2015. DOI: 10.1038/nature14236

[9] *Intelligent Transport Systems (ITS): Congestion Control Mechanisms for the C-V2X PC5 interface: Access layer part*, document TS 103 574 V1.1.1, Nov. 2018.

[10] *On-Board System Requirements for V2V Safety Communications*, Standard SAE J2945/1, Mar. 2016.

BIOGRAPHY

Byeong Cheol Jeon (Member)



2022 : BS degree in Electronic Engineering, Hanbat National University.

2022~ : MS degree in Electronic Engineering, Hanbat National University.

Woo Yeol Yang (Member)



2021 : BS degree in Electronic Engineering, Hanbat National University.

2023 : MS degree in Electronic Engineering, Hanbat National University.

Han-Shin Jo (Member)



2001 : BS degree in Electrical and Electronics Engineering, Yonsei University

2004 : MS degree in Electrical and Electronics Engineering, Yonsei University

2009 : Ph.D degree in Electrical and Electronics Engineering, Yonsei University

2009~2011 : Postdoctoral researcher of the Department of Electrical and Computer Engineering, the University of Texas at Austin

2011~2012 : Senior researcher, Samsung Electronics

2012~2023 : Prof. of Dept. of Electronic Engineering, Hanbat National University

2023~ : Prof. of Dept. of Automotive Engineering, Hanyang University